SPECIFICATION

Electronic Version 1.2.8

Stylesheet Version 1.0

[Method For Click-Stream Analysis Using Web Directory Reverse Categorization]

Cross Reference to Related Applications

This application claims priority under 35 U.S.C. § 119(e) to United States Patent Application No. 60/208,889 entitled "Method For Click-Stream Analysis Using Web Directory Reverse Categorization" by McCullough filed June 2, 2000, and United States Patent Application No. 60/208,890 entitled "Method For Continuous, Frame-Specific Click-Stream Recording" by McCullough June 2, 2000. This is also related to United States Patent Application No. (Attorney Docket VIGN1250-1) entitled "Method For Continuous, Frame-Specific Click-Stream Recording" by McCullough filed of even date. All applications listed in this paragraph and are assigned to the current assignee hereof and are incorporated herein by reference.

Background of Invention

[0001]   *FIELD OF THE INVENTION*

[0002]   This invention relates in general to methods and data processing system readable media, and more particularly, to methods of profiling a user and data processing system readable media for carrying out those methods.

[0003]   *DESCRIPTION OF THE RELATED ART*

[0004]

Many different Internet Service Providers ("ISPs") or sites are interested in creating or using profiles of their users. One way to obtain a profile for a user is to rely on information that the user may voluntarily provide. One problem with this is

that the user may lie or give inaccurate, misleading, or incomplete data regarding himself or herself.

[0005]   Another attempt at establishing a profile for a user may include reviewing the Internet Protocol ("IP") address of the user and determining the geographic location and time zone for the user. However, that information may be misleading depending upon the ISP and where it is located. For example, if the user is using America Online, the user may be listed as residing in Virginia and being in the Eastern Time zone. In reality, the user may be in a different country and residing within a time zone many hours different from the Eastern Time zone. Additionally, many different users may be using the same computer (same IP address). Clearly, such information is of little use in determining interests or other information regarding an individual. Even if the information regarding state and time zone are correct and only one user is using a specific IP address with a local ISP, this may yield little useful information about the individual user. The ISP may have thousands or millions of users.

Summary of Invention

[0006]   A user profile can be generated using a table that tracks a user by category information and temporal information. In one embodiment, the user profile can be generated with relatively minimal resources and relies on the user behavior rather than information provided by the user or just the IP address. In one configuration, a network access provider or network site may track a user's network addresses by timestamp and correlate the network addresses to meta tags for the network address that are obtained for free or using a third party source. A table including a user identifier, corresponding category, and timestamp may provide a basis for forming a user profile.

[0007]
     In one set of embodiments, a method of profiling a user can comprise accessing first data including a first identifier for the user, network addresses accessed by the user, and temporal information related to the user identifier and the network addresses. The method can also comprise accessing second data including at least some the network addresses and corresponding category

information for each of those network addresses. The method can further comprise generating a user profile based at least in part on the first identifier, category information, and at least some of the temporal information.

[0008]   In a specific embodiment, generating the user profile can include creating a table (third data), comparing the data for the user within the table to existing profiles, and associating the user with a particular profile. In another embodiment, the method can further comprise selecting marketing information regarding an item, where the selection is based at least in part on the user profile, and sending the marketing information to the user. In still another embodiment, the user profile can be sold.

[0009]   In other embodiments, a data processing system readable medium can have code embodied within it. The code can include instructions executable by a data processing system. The instructions may be configured to cause the data processing system to perform the methods described herein.

[0010]   The foregoing general description and the following detailed description are exemplary and explanatory only are not restrictive of the invention, as claimed.

Brief Description of Drawings

[0011]   The present invention is illustrated by way of example and not limitation in the accompanying figures, in which:

[0012]   FIG. 1 includes an illustration of a hardware architecture for carrying out methods of profiling a user;

[0013]   FIG. 2 includes an illustration of a data processing system readable medium including software code; and

[0014]   FIG. 3 includes a process flow diagram illustrating generation and use of a user profile.

[0015]

Skilled artisans appreciate that elements in the figures are illustrated for simplicity and clarity and have not necessarily been drawn to scale. For example,

the dimensions of some of the elements in the figures may be exaggerated relative to other elements to help to improve understanding of embodiments of the present invention.

Detailed Description

[0016]   Reference is now made in detail to the exemplary embodiments of the invention, examples of which are illustrated in the accompanying drawings. Wherever possible, the same reference numbers will be used throughout the drawings to refer to the same or like parts (elements).

[0017]   A method can be used to profile a user using network addresses and temporal information of the network addresses when the user requests or is sent information for those network addresses. The method can comprise accessing first data including a user identifier, network addresses accessed by the user, and temporal information. The method can also comprise accessing second data including at least some of the network addresses and corresponding category information for each of those network addresses. The method can further comprise generating the user profile based at least in part on the user identifier, category information, and at least some of the temporal information.

[0018]   A few terms are defined or clarified to aid in understanding the descriptions that follow. A network includes an interconnected set of server and client computers over a publicly available medium (e.g., the internet) or over an internal (company-owned) system. A user at a client computer may gain access to the network using a network access provider. An Internet Service Provider ("ISP") is a common type of network access provider. A network address includes information that can be used by a server computer to locate information, whether internal to that server computer or at a different, remote computer or database. URLs are examples of network addresses. A network site corresponds to a location specified by a network address. A web site is a common type of network site. Note that the examples given within this paragraph are for purposes of illustration and not limitation.

[0019]   Before discussing details of embodiments of present invention, a hardware architecture for using embodiments is described. FIG. 1 illustrates an exemplary architecture and includes a client computer 12 that is bi-directionally coupled to a network 14, and a server computer 16 that is bi-directionally coupled to the network 14 and database 18. The client computer 12 includes a central processing unit ("CPU") 120, a read-only memory ("ROM") 122, a random access memory ("RAM") 124, a hard drive ("HD") or storage memory 126, and input/output device (s) ("I/O") 128. The I/O devices 128 can include a keyboard, monitor, printer, electronic pointing device (e.g., mouse, trackball, etc.), or the like. The server computer 16 can include a CPU 160, ROM 162, RAM 164, HD 166, and I/O 168.

[0020]   Each of the client computer 12 and the server computer 16 is an example of a data processing system. ROM 122 and 162, RAM 124 and 164, HD 126 and 166, and the database 18 include media that can be read by the CPU 120 or 160. Therefore, each of these types of memories includes a data processing system readable medium. These memories may be internal or external to the computers 12 and 16.

[0021]   The methods described herein may be implemented in suitable software code that can reside within ROM 122 or 162, RAM 124 or 164, or HD 126 or 166. FIG. 2 illustrates a combination of software code elements 204, 206, and 208 that are embodied within a data processing system readable medium 202 on HD 166. In addition to those types of memories previously described, the instructions in an embodiment of the present invention may be contained on a data storage device with a different data processing system readable storage medium. The instructions may be stored as software code elements on a DASD array, magnetic tape, floppy diskette, electronic read-only memory, optical storage device, CD ROM or other appropriate data processing system readable medium or storage device.

[0022]

In an illustrative embodiment of the invention, the computer-executable instructions may be lines of compiled C^{++}, Java, or other language code. Other architectures may be used. For example, the functions of the client computer 12 may be incorporated into the server computer 16, and vice versa. In alternative

embodiments, more server computers can be used, such that the functions of server computer 16 are performed by a combination of server computers. FIG. 3 includes an illustration, in the form of a flow diagram, of the structure of such a software program.

[0023]   Communications between the client computer 12 and the server computer 16 can be accomplished using radio frequency, electronic, or optical signals. When a user (human) is at the client computer 12, the client computer 12 may convert the signals to a human understandable form when sending a communication to the user and may convert input from the user to appropriate signals to be used by the client computer 12 or the server computer 16.

[0024]   Attention is now directed to methods of profiling users as illustrated in FIG. 3. The method can comprise accessing data in a first table that includes user identifiers, network addresses, and temporal information (block 302). The first table is typically part of the database 18. The first table may be organized in rows and columns. The columns may correspond to user identifiers, network addresses, temporal information, and the like. The rows may include an entry for each time a network address is requested by users. Note that the network addresses may be for different network domains of different companies that do not significantly own or control one another (i.e., no more than approximately 10% ownership of a company or voting shares of the company).

[0025]   In one embodiment, a user can "surf" the network on client computer 12, while the server computer 16 collects and stores the user identifier, network addresses, and timestamps within the first table in the database 18. The user may be routed to other server computers (not shown) that correspond to the network sites requested by the user at client computer 12. The server computer 16 may access and read the information within the first table within database 18. Other users (not shown) may be using the network 14 and server computer 16.

[0026]   The method can also comprise accessing data in a second table that includes network addresses and corresponding category information (304). This act may be performed by the server computer 16 in one embodiment. The data for the second

table may reside in the storage device 18 or, alternatively, may reside within a database at an external location of a third party (not shown). For example, the second table may be part of the Netscape Open Directory Project or may be provided by LookSmart, Ltd. of San Francisco, California. Either of these third-party sources may provide a categorization of each Internet site. The categorization may be in the form of meta tags, which are metadata that correspond to the network addresses. Skilled artisans may realize that meta tags are used by search engines, such as Yahoo!, Lycos, Excite, and the like in performing user specified searches for information. To the inventor's knowledge, the meta tags for the search engines have not been used in generating user profiles. A short text portion of a URL may optionally be used. In any event, a table of network addresses and categories can be created or accessed using a third-party source that is different from the user and the party that significantly owns or controls the server computer 16.

[0027]   The method can further include creating a third table that includes user identifiers, category information, and temporal information as shown in block 322. The table can be created by matching the network addresses in the first and second tables. Other information may be included within the third table. In one embodiment, the third table can sort by user identifier and time to show categories in chronological or reverse chronological order.

[0028]   Different method may be used to generate the user profile. The methods described herein are only to illustrate and not limit the present invention. In one embodiment, the information in the third table can be used in generating a user profile by performing click-stream analysis. To generate the user profile, the data within the third table can be compared to existing profiles (block 342). The existing profiles can represent another person or can represent a theoretical individual having attributes similar to a person. For each user, his or her information from the third table can be compared to the existing profiles. The method can also include associating each of the users with a profile that most closely matches that particular user (block 344). Alternatively, data mining rules may be used to determine the characteristics of the user, and the user profile may include a set of discrete, but potentially interrelated, characteristics.

[0029]   The information gathered may be used in wide variety of ways. The examples described below are meant to illustrate and not limit the present invention. In one embodiment, the information may be useful for providing information or offering an item to the user. As used herein, "item" means a product or a service. The method can include selecting marketing information regarding an item based on the user profiles (block 362). The marketing information may vary widely from embodiment to embodiment. In one embodiment, the marketing information may include banner advertisements that may be displayed near a periphery of a view as seen by the user on I/O 128 of client computer 12. A network access provider may perform the selection of the banner advertisements. Alternatively, network sites may be owned or controlled by a company that sells the item. In this example, the marketing information may comprise an offer to sell the item to the user.

[0030]   The method can further comprise applying one or more filters as shown in block 364. These filters can be used to remove information that may have been previously seen by the user, to remove information regarding some items having lower profit margins from a collection of items, to remove information that may be offensive or objectionable to be user, or the like. The method can further comprise an optional act of sending the filtering marketing information to the user as shown in block 366.

[0031]   In another embodiment, a company creating user profiles using data within the tables may sell the user profiles to a marketing company, a mailing service, or other third party. Alternatively, the information can be used in performing internal analytics within the company. After reading this specification, skilled artisans appreciate that the number and variety of uses of the user profiles is nearly limitless.

[0032]

Embodiments of the present invention may allow creation of a user profile without having to rely on the user for his or her input or having to rely upon information only within a HyperText Transfer Protocol (HTTP) stream, such as an IP address. The click stream may be the only source of user information used for the profile. Solely by a using a user's chronological click stream and network address

categorization, a reasonably accurate user profile can be obtained for an individual at a reasonably low cost without complex software code requirements.

[0033]   In the foregoing specification, the invention has been described with reference to specific embodiments. However, one of ordinary skill in the art appreciates that various modifications and changes can be made without departing from the scope of the present invention as set forth in the claims below. Accordingly, the specification and figures are to be regarded in an illustrative rather than a restrictive sense, and all such modifications are intended to be included within the scope of present invention.

[0034]   Benefits, other advantages, and solutions to problems have been described above with regard to specific embodiments. However, the benefits, advantages, solutions to problems, and any element(s) that may cause any benefit, advantage, or solution to occur or become more pronounced are not to be construed as a critical, required, or essential feature or element of any or all the claims. As used herein, the terms "comprises," "comprising," or any other variation thereof, are intended to cover a non-exclusive inclusion, such that a process, method, article, or apparatus that comprises a list of elements does not include only those elements but may include other elements not expressly listed or inherent to such process, method, article, or apparatus.